10/561867

IAP20 Rec'd PCT/PTO 21 DEC 2005

PROTEIN CRYSTAL COMPRISING THE PROCESSIVITY CLAMP FACTOR OF DNA POLYMERASE AND A LIGAND, AND ITS USES

The present invention relates to a protein crystal comprising the processivity clamp factor of DNA polymerase and a peptide comprising all or part of the processivity clamp factor binding sequence of a processivity clamp factor interacting protein, and its uses, in particular for the screening, the design or the modification of ligands of the processivity clamp factor of DNA polymerase.

The presence of lesions on DNA may severely impair its replication and have dramatic consequences on cells survival. Beside the activity of efficient repair processes, which remove most of the lesions from DNA before replication occurs, the replisome is able to cope with replication blocking DNA lesions, thanks to specialized biochemical processes refered to as damaged DNA tolerance pathways. Translesion synthesis (TLS) is one of these mechanisms which requires the incorporation of a nucleotide opposite and past the lesion. Depending on the nature of the incorporated nucleotide relative to the parental sequence, the TLS process is error-free or mutagenic. TLS has recently gained much understanding, with the discovery of specialized DNA polymerases, which are able to replicate through lesions which otherwise impede the progression of DNA polymerases involved in replication. These new polymerases have been found in both prokaryotes and eukaryotes and most of them have been classified in the Y superfamily (Ohmori *et al.*, 2001). In *Escherichia coli*, two such polymerases have been identified, Pol IV (DinB) (Wagner *et al.*, 1999) and Pol V (Tang et al., 1999; Reuven *et al.*, 1999), whereas Pol II polymerase has also been shown to perform TLS, although it belongs to the B family (Napolitano *et al.*, 2000; Becherel *et al.*, 2001; Fuchs *et al.*, 2001). Interestingly, all these three polymerase genes are part of the SOS network and are induced upon the arrest of replication due to the presence of replicase blocking lesions onto DNA.

The discovery of translesional polymerases (Ohmori *et al.*, 2001) resulted in a major modification of the molecular model of TLS and resulting lesion induced mutagenesis. The previous model, essentially built on genetic experiments in *E. coli* (Bridges and Woodgates, 1985) suggested that the replicative polymerase stalled at blocking lesions was assisted by SOS induced proteins, whose functions were expected to facilitate the polymerase progression through the lesion by increasing its anchoring

BEST AVAILABLE COPY

CONFIRMATION COPY

onto modified DNA or by reducing its fidelity either by alteration of the correct nucleotide selection process and/or by inhibition of its proofreading activity. The current new model (Cordonnier *et al.*, 1999) proposes that the blocked replicative polymerase is replaced by one or several TLS polymerases that cooperate at different steps of the translesional process, namely incorporation opposite the lesion and elongation of the lesion terminus, to ensure an efficient bypass of the lesion. These polymerases further dissociate from the DNA substrate and the replicative enzyme resumes its synthesis function.

It was demonstrated that prokaryotic and eukaryotic replicative polymerases (Pol III holoenzyme of *E. coli*, pol C, eukaryotic pol δ and pol ε) physically interact with their respective processivity clamp factor, also called sliding clamp. Moreover, all prokaryotic and most eukaryotic TLS polymerases also interact with their processivity clamp factor (Lenne-Samuel *et al.*, 2002; Wagner *et al.*, 2000; Becherel *et al.*, 2002; Haracska *et al.*, 2002; Haracska *et al.*, 2001a; Haracska *et al.*, 2001b). These clamps, which act by increasing the replicative polymerase processivity (Bruck and O'Donnel, 2001), are homodimeric (β of *E. coli*) or homotrimeric (gp45 of T4/RB69 or PCNA in eukaryotes) toroid-shape molecules that are loaded onto DNA near primer-template junctions, by specific clamp loader complexes (e.g. the so-called γ complex in *E. coli* and RFC in eukaryotes). The β and PCNA monomers fold into structurally similar subdomains (3 and 2, respectively), despite a lack of internal homology in their amino acids sequences, so that the ring presents a pseudo-six-fold symetry. A consensus pentapeptidic sequence, QL(SD)LF, conserved among eubacteria, was identified in most of the β-binding proteins as the motif mediating their connection with the clamp, through hydrophobic interactions (Dalrymple *et al.*, 2001). Similarly, a eukaryotic PCNA (or alternative sliding clamps) consensus binding sequence has been identified. A recent study in *E. coli* demonstrated that the integrity of this motif is absolutely required for the inducible polymerases to perform TLS: Pol IV and Pol II mutant proteins deleted for their β-clamp binding motif retain their polymerase activity, but loose their functions in the TLS process *in vivo*, highlightening the fact that their functional interaction with β is crucial for translesion DNA synthesis and mutagenesis (Becherel *et al.*, 2002; Lenne-Samuel *et al.*, 2002).

The presence of several TLS polymerases within a single organism has remained a puzzling question. Analysis of the TLS process in *E. coli* indicated that, depending on

both the nature of the lesion and the local DNA sequence, one or several TLS polymerases may participate to a single TLS event (Napolitano *et al.*, 2000; Wagner *et al.*, 2002). TLS appears as a complex process where a pool of low fidelity polymerases replace the highly stringent replisome and eventually exchange mutually to accommodate the large variety of DNA lesions and to ensure ultimately the completion of DNA replication. Whether this polymerase switching process is somehow coordinated or simply occurs on the basis of competition between the different TLS polymerases is not yet known.

An object of the invention is to provide a method to obtain ligands of the processivity clamp factor which would impair the interaction between the sliding clamp and its interacting proteins.

Such ligands might be useful for the preparation of drugs for the treatment of bacterial diseases or of proliferative disorders.

The invention follows on from the solving by the Inventors of the structure of a co-cristal obtained between the β clamp of *E. coli* and the 16 residues C-terminal peptide of Pol IV DNA polymerase (P16) of *E. coli* containing its β-binding sequence, from the identification of the peptide binding site on β and from the description of the interactions between P16 and β residues.

The Invention also follows on from the results of experiments carried out by the Inventors showing that P16 competes with Pol IV, but also with the α subunit of the *E. coli* replicative Pol III holoenzyme, for binding to β, thus inhibiting their β dependent polymerase activity.

The present invention relates to a protein crystal comprising the processivity clamp factor of DNA polymerase and a peptide of about 3 to about 30 amino acids, in particular of about 16 amino acids, said peptide comprising all or part of the processivity clamp-factor binding sequence of a processivity clamp factor-interacting protein, such as prokaryotic Pol I, Pol II, Pol III, Pol IV, Pol V, MutS, ligase I, α subunit of DNA polymerase, UmuD or UmuD', or eukaryotic pol ε, pol δ, pol η, pol ι, pol κ.

Other processivity clamp factor interacting proteins are notably described in Tsurimoto *et al.* (1999).

The expression "processivity clamp factor of DNA polymerase" refers to *dnaN* genes products and their functional analogs in prokaryotes, and *PCNA* genes products

and their functional analogs and orthologs in eukaryotes. It can also be referred to as a sliding clamp. It is notably described in Kong *et al.* (1992) and Gulbis *et al.* (1996).

"Pol I", "Pol II", "Pol III", "Pol IV", "Pol V" respectively refer to DNA polymerase I, II, III, IV and V, in bacteria, such as *E. coli*, as reviewed in Friedberg *et al.* (2000a), and Friedberg et al. (2000b).

"MutS" refers to the product of the *mutS* gene in *E. coli*, and functional analogs and orthologs thereof, involved in mismatch repair.

"Ligase I" refers to the product of the *lig* gene in *E. coli*, and functional analogs and orthologs thereof.

"α subunit of DNA polymerase" refers to the product of the *dnaE* gene in *E. coli*, and functional analogs and orthologs thereof.

"UmuD" refers to the product of the *umuD* gene in *E. coli*, and functional analogs and orthologs thereof.

"Pol ε", "pol δ", "pol η", "pol ι", "pol κ" refer to eukaryotic polymerases as reviewed in Friedberg *et al.* (2000a), and Friedberg et al. (2000b).

The invention more particularly relates to a protein crystal as defined above, wherein the processivity clamp factor of DNA polymerase is the β subunit of DNA polymerase, in particular the β subunit of DNA polymerase III of *Escherichia coli*, and the peptide has the following sequence:

VTLLDPQMERQLVLGL (SEQ ID NO: 1)

The β subunit of DNA polymerase III of *Escherichia coli* is in particular described in Kong *et al.* (1992).

The invention more particularly relates to a protein crystal as defined above, comprising the β subunit of DNA polymerase III of *Escherichia coli* and the peptide of SEQ ID NO: 1, said crystal belonging to the triclinic space group P1 and its cell dimensions being approximately a = 41.23 Å, b = 65.22 Å, c = 73.38 Å, $\alpha = 73.11°$, $\beta = 85.58°$, $\gamma = 85.80°$.

The expression "triclinic space goup P1" refers to a nomenclature well known to the man skilled in the art, it is in particular described in "International tables for X-ray crystallography", Vol. 1 (The Kynoch press, Birmingham, England, 1968)

The expression "cell dimensions" refers to the geometrical description of the smallest volume being repeated in the three dimensions to build the crystal.

The invention more particularly relates to a protein crystal as defined above, characterized by the atomic coordinates such as obtained by the X-ray diffraction of said crystal, said atomic coordinates being represented in Figure 1.

The expression "atomic coordinates" refers to the three coordinates X, Y, Z (given in Å, 1Å=10^{-10} m) necessary to describe the exact position of each atom in the molecule.

The expression "X-ray diffraction" refers to the phenomenon following which X-rays are scattered in a specific way by a crystal.

Two major X-ray sources can be used: a rotating anode, which is a usual laboratory equipment and/or a synchrotron which is a large-scale equipment, such as the European Synchrotron Radiation Facility (ESRF) in Grenoble, France.

The general methodology to obtain atomic coordinates from X-ray diffraction of a crystal is well known to man skilled in the art, briefly it consists in measuring the intensities of the numerous secondary X-rays beams resulting from the diffraction by the crystal of an incident X-ray beam.

The invention more particularly relates to a protein crystal as defined above, characterized by the atomic coordinates representing the peptide and the peptide binding site of the β subunit of DNA polymerase III of *Escherichia coli*, and being as follows:

ATOM	4045	N	LEU	B	155	5.874	17.816	22.109	1.00	1.00	B
ATOM	4046	CA	LEU	B	155	6.029	16.359	22.087	1.00	1.00	B
ATOM	4047	CB	LEU	B	155	5.055	15.686	23.064	1.00	1.00	B
ATOM	4048	CG	LEU	B	155	5.260	16.046	24.536	1.00	1.00	B
ATOM	4049	CD1	LEU	B	155	4.256	15.237	25.360	1.00	1.00	B
ATOM	4050	CD2	LEU	B	155	6.686	15.757	24.980	1.00	1.00	B
ATOM	4051	C	LEU	B	155	5.808	15.776	20.682	1.00	1.00	B
ATOM	4052	O	LEU	B	155	6.177	14.613	20.431	1.00	1.00	B
ATOM	4177	N	THR	B	172	9.112	11.246	22.902	1.00	1.00	B
ATOM	4178	CA	THR	B	172	8.212	10.730	23.917	1.00	1.00	B
ATOM	4179	CB	THR	B	172	8.776	11.014	25.344	1.00	1.00	B
ATOM	4180	OG1	THR	B	172	7.931	10.400	26.328	1.00	1.00	B
ATOM	4181	CG2	THR	B	172	8.870	12.532	25.619	1.00	1.00	B
ATOM	4182	C	THR	B	172	6.805	11.269	23.709	1.00	1.00	B
ATOM	4183	O	THR	B	172	6.588	12.352	23.145	1.00	1.00	B
ATOM	4192	N	GLY	B	174	4.562	10.770	26.397	1.00	1.00	B
ATOM	4193	CA	GLY	B	174	3.992	10.745	27.737	1.00	1.00	B
ATOM	4194	C	GLY	B	174	3.762	9.337	28.266	1.00	1.00	B
ATOM	4195	O	GLY	B	174	3.667	9.141	29.489	1.00	1.00	B
ATOM	4196	N	HIS	B	175	3.650	8.349	27.375	1.00	1.00	B
ATOM	4197	CA	HIS	B	175	3.440	-6.953	27.796	1.00	1.00	B
ATOM	4198	CB	HIS	B	175	2.313	6.309	26.977	1.00	1.00	B
ATOM	4199	CG	HIS	B	175	0.992	6.997	27.119	1.00	1.00	B
ATOM	4200	CD2	HIS	B	175	0.106	7.435	26.193	1.00	1.00	B
ATOM	4201	ND1	HIS	B	175	0.420	7.255	28.345	1.00	1.00	B
ATOM	4202	CE1	HIS	B	175	-0.763	7.817	28.170	1.00	1.00	B
ATOM	4203	NE2	HIS	B	175	-0.977	7.938	26.875	1.00	1.00	B
ATOM	4204	C	HIS	B	175	4.706	6.135	27.641	1.00	1.00	B
ATOM	4205	O	HIS	B	175	4.990	5.212	28.403	1.00	1.00	B
ATOM	4207	CA	ARG	B	176	6.711	5.768	26.422	1.00	18.30	B
ATOM	4208	CB	ARG	B	176	6.575	4.633	25.398	1.00	19.53	B
ATOM	4209	CG	ARG	B	176	6.329	5.094	23.954	1.00	22.88	B
ATOM	4210	CD	ARG	B	176	4.876	4.888	23.657	1.00	22.11	B
ATOM	4211	NE	ARG	B	176	4.435	5.312	22.314	1.00	22.09	B
ATOM	4212	CZ	ARG	B	176	4.555	4.591	21.202	1.00	20.17	B
ATOM	4213	NH1	ARG	B	176	5.159	3.403	21.213	1.00	17.04	B
ATOM	4214	NH2	ARG	B	176	3.914	4.977	20.120	1.00	20.02	B
ATOM	4215	C	ARG	B	176	7.684	6.807	25.902	1.00	17.30	B

ATOM	4216	O	ARG	B	176	7.255	7.860	25.374	1.00	18.10	B
ATOM	4217	N	LEU	B	177	8.957	6.504	26.080	1.00	17.97	B
ATOM	4218	CA	LEU	B	177	10.049	7.360	25.633	1.00	17.85	B
ATOM	4219	CB	LEU	B	177	10.664	8.095	26.827	1.00	18.29	B
ATOM	4220	CG	LEU	B	177	11.921	8.955	26.611	1.00	16.28	B
ATOM	4221	CD1	LEU	B	177	11.819	10.163	27.559	1.00	19.52	B
ATOM	4222	CD2	LEU	B	177	13.191	8.172	26.839	1.00	19.12	B
ATOM	4223	C	LEU	B	177	11.110	6.517	24.964	1.00	18.45	B
ATOM	4224	O	LEU	B	177	11.291	5.329	25.281	1.00	18.33	B
ATOM	4710	N	PRO	B	242	11.254	17.279	27.890	1.00	1.00	B
ATOM	4711	CD	PRO	B	242	9.987	16.826	27.286	1.00	1.00	B
ATOM	4712	CA	PRO	B	242	11.660	16.404	28.997	1.00	1.00	B
ATOM	4713	CB	PRO	B	242	10.688	15.230	28.874	1.00	1.00	B
ATOM	4714	CG	PRO	B	242	9.448	15.869	28.336	1.00	1.00	B
ATOM	4715	C	PRO	B	242	13.124	15.947	28.987	1.00	1.00	B
ATOM	4716	O	PRO	B	242	13.728	15.748	27.925	1.00	1.00	B
ATOM	4748	N	ARG	B	246	16.133	11.840	33.560	1.00	1.00	B
ATOM	4749	CA	ARG	B	246	15.239	11.808	34.707	1.00	1.00	B
ATOM	4750	CB	ARG	B	246	14.755	13.227	34.984	1.00	1.00	B
ATOM	4751	CG	ARG	B	246	15.880	14.252	35.113	1.00	1.00	B
ATOM	4752	CD	ARG	B	246	16.443	14.295	36.529	1.00	1.00	B
ATOM	4753	NE	ARG	B	246	15.374	14.318	37.524	1.00	1.00	B
ATOM	4754	CZ	ARG	B	246	14.316	15.126	37.477	1.00	1.00	B
ATOM	4755	NH1	ARG	B	246	14.169	15.992	36.481	1.00	1.00	B
ATOM	4756	NH2	ARG	B	246	13.396	15.067	38.430	1.00	1.00	B
ATOM	4757	C	ARG	B	246	14.022	10.889	34.566	1.00	1.00	B
ATOM	4758	O	ARG	B	246	13.384	10.536	35.560	1.00	1.00	B
ATOM	4759	N	VAL	B	247	13.695	10.532	33.327	1.00	1.00	B
ATOM	4760	CA	VAL	B	247	12.553	9.675	33.018	1.00	1.00	B
ATOM	4761	CB	VAL	B	247	12.061	9.942	31.585	1.00	1.00	B
ATOM	4762	CG1	VAL	B	247	10.930	8.991	31.216	1.00	1.00	B
ATOM	4763	CG2	VAL	B	247	11.624	11.391	31.462	1.00	1.00	B
ATOM	4764	C	VAL	B	247	12.962	8.218	33.133	1.00	1.00	B
ATOM	4765	O	VAL	B	247	12.125	7.334	33.308	1.00	1.00	B
ATOM	4996	N	PHE	B	278	-7.702	-1.352	24.244	1.00	1.00	B
ATOM	4997	CA	PHE	B	278	-6.698	-1.155	25.300	1.00	1.00	B
ATOM	4998	CB	PHE	B	278	-7.318	-1.432	26.663	1.00	1.00	B
ATOM	4999	CG	PHE	B	278	-8.431	-0.459	27.021	1.00	1.00	B
ATOM	5000	CD1	PHE	B	278	-8.142	0.882	27.268	1.00	1.00	B
ATOM	5001	CD2	PHE	B	278	-9.760	-0.869	27.021	1.00	1.00	B
ATOM	5002	CE1	PHE	B	278	-9.177	1.816	27.508	1.00	1.00	B
ATOM	5003	CE2	PHE	B	278	-10.795	0.052	27.258	1.00	1.00	B
ATOM	5004	CZ	PHE	B	278	-10.496	1.391	27.500	1.00	1.00	B
ATOM	5005	C	PHE	B	278	-5.403	-1.957	25.131	1.00	1.00	B
ATOM	5006	O	PHE	B	278	-4.356	-1.582	25.677	1.00	1.00	B
ATOM	5332	N	ASN	B	320	0.635	-2.143	27.431	1.00	1.00	B
ATOM	5333	CA	ASN	B	320	-0.051	-1.983	26.158	1.00	1.00	B
ATOM	5334	CB	ASN	B	320	-0.055	-0.504	25.796	1.00	1.00	B
ATOM	5335	CG	ASN	B	320	-0.561	-0.259	24.407	1.00	1.00	B
ATOM	5336	OD1	ASN	B	320	-0.226	-0.997	23.481	1.00	1.00	B
ATOM	5337	ND2	ASN	B	320	-1.362	0.791	24.242	1.00	1.00	B
ATOM	5338	C	ASN	B	320	0.927	-2.745	25.249	1.00	1.00	B
ATOM	5339	O	ASN	B	320	2.093	-2.350	25.102	1.00	1.00	B
ATOM	5353	N	TYR	B	323	2.932	-0.853	22.482	1.00	1.00	B
ATOM	5354	CA	TYR	B	323	4.110	-0.088	22.908	1.00	1.00	B
ATOM	5355	CB	TYR	B	323	3.878	0.590	24.259	1.00	1.00	B
ATOM	5356	CG	TYR	B	323	2.813	1.668	24.294	1.00	1.00	B
ATOM	5357	CD1	TYR	B	323	2.397	2.314	23.127	1.00	1.00	B
ATOM	5358	CE1	TYR	B	323	1.458	3.374	23.170	1.00	1.00	B
ATOM	5359	CD2	TYR	B	323	2.284	2.093	25.509	1.00	1.00	B
ATOM	5360	CE2	TYR	B	323	1.354	3.166	25.567	1.00	1.00	B
ATOM	5361	CZ	TYR	B	323	0.957	3.790	24.399	1.00	1.00	B
ATOM	5362	OH	TYR	B	323	0.112	4.886	24.453	1.00	1.00	B
ATOM	5363	C	TYR	B	323	5.327	-1.018	23.041	1.00	1.00	B
ATOM	5364	O	TYR	B	323	6.468	-0.646	22.726	1.00	1.00	B
ATOM	5519	N	VAL	B	344	3.837	-1.100	39.291	1.00	1.00	B
ATOM	5520	CA	VAL	B	344	3.324	0.227	39.030	1.00	1.00	B
ATOM	5521	CB	VAL	B	344	2.676	0.818	40.318	1.00	1.00	B
ATOM	5522	CG1	VAL	B	344	1.474	-0.026	40.725	1.00	1.00	B
ATOM	5523	CG2	VAL	B	344	3.687	0.847	41.456	1.00	1.00	B
ATOM	5524	C	VAL	B	344	4.405	1.163	38.512	1.00	1.00	B
ATOM	5525	O	VAL	B	344	4.199	2.365	38.405	1.00	1.00	B
ATOM	5532	N	SER	B	346	7.618	2.153	35.615	1.00	21.53	B
ATOM	5533	CA	SER	B	346	8.060	2.002	34.239	1.00	21.50	B
ATOM	5534	CB	SER	B	346	8.655	3.320	33.722	1.00	21.47	B
ATOM	5535	OG	SER	B	346	9.793	3.703	34.474	1.00	26.08	B
ATOM	5536	C	SER	B	346	9.107	0.914	34.106	1.00	20.70	B

```
ATOM   5537  O   SER B 346       9.755   0.521  35.078  1.00 21.55      B
ATOM   5632  N   VAL B 360      11.730   3.546  27.545  1.00  1.00      B
ATOM   5633  CA  VAL B 360      11.023   3.501  28.812  1.00  1.00      B
ATOM   5634  CB  VAL B 360      11.276   4.794  29.641  1.00  1.00      B
ATOM   5635  CG1 VAL B 360      10.448   4.742  30.934  1.00  1.00      B
ATOM   5636  CG2 VAL B 360      12.753   4.923  29.937  1.00  1.00      B
ATOM   5637  C   VAL B 360       9.562   3.381  28.501  1.00  1.00      B
ATOM   5638  O   VAL B 360       9.008   4.188  27.753  1.00  1.00      B
ATOM   5639  N   VAL B 361       8.905   2.372  29.069  1.00 19.72      B
ATOM   5640  CA  VAL B 361       7.488   2.188  28.831  1.00 18.92      B
ATOM   5641  CB  VAL B 361       7.216   0.872  28.069  1.00 18.99      B
ATOM   5642  CG1 VAL B 361       5.743   0.769  27.716  1.00 18.31      B
ATOM   5643  CG2 VAL B 361       8.065   0.839  26.786  1.00 17.76      B
ATOM   5644  C   VAL B 361       6.793   2.100  30.167  1.00 19.47      B
ATOM   5645  O   VAL B 361       7.232   1.362  31.038  1.00 16.90      B
ATOM   5646  N   MET B 362       5.737   2.885  30.318  1.00  1.00      B
ATOM   5647  CA  MET B 362       4.962   2.882  31.540  1.00  1.00      B
ATOM   5648  CB  MET B 362       4.226   4.206  31.682  1.00  1.00      B
ATOM   5649  CG  MET B 362       3.918   4.589  33.122  1.00  1.00      B
ATOM   5650  SD  MET B 362       5.405   4.806  34.163  1.00  1.00      B
ATOM   5651  CE  MET B 362       4.575   4.880  35.731  1.00  1.00      B
ATOM   5652  C   MET B 362       3.949   1.731  31.471  1.00  1.00      B
ATOM   5653  O   MET B 362       3.385   1.438  30.410  1.00  1.00      B
ATOM   5654  N   PRO B 363       3.698   1.069  32.599  1.00  1.00      B
ATOM   5655  CD  PRO B 363       4.521   1.025  33.818  1.00  1.00      B
ATOM   5656  CA  PRO B 363       2.729  -0.038  32.579  1.00  1.00      B
ATOM   5657  CB  PRO B 363       3.155  -0.883  33.776  1.00  1.00      B
ATOM   5658  CG  PRO B 363       3.665   0.160  34.754  1.00  1.00      B
ATOM   5659  C   PRO B 363       1.272   0.395  32.672  1.00  1.00      B
ATOM   5660  O   PRO B 363       0.959   1.574  32.811  1.00  1.00      B
ATOM   5661  N   MET B 364       0.368  -0.568  32.537  1.00  1.00      B
ATOM   5662  CA  MET B 364      -1.037  -0.272  32.674  1.00  1.00      B
ATOM   5663  CB  MET B 364      -1.780  -0.391  31.332  1.00  1.00      B
ATOM   5664  CG  MET B 364      -1.636  -1.670  30.568  1.00  1.00      B
ATOM   5665  SD  MET B 364      -2.386  -1.510  28.872  1.00  1.00      B
ATOM   5666  CE  MET B 364      -4.155  -1.253  29.308  1.00  1.00      B
ATOM   5667  C   MET B 364      -1.602  -1.218  33.725  1.00  1.00      B
ATOM   5668  O   MET B 364      -0.999  -2.251  34.035  1.00  1.00      B
ATOM   5669  N   ARG B 365      -2.732  -0.836  34.307  1.00  1.00      B
ATOM   5670  CA  ARG B 365      -3.383  -1.655  35.324  1.00  1.00      B
ATOM   5671  CB  ARG B 365      -4.029  -0.756  36.394  1.00  1.00      B
ATOM   5672  CG  ARG B 365      -4.785  -1.490  37.505  1.00  1.00      B
ATOM   5673  CD  ARG B 365      -3.859  -2.316  38.398  1.00  1.00      B
ATOM   5674  NE  ARG B 365      -4.571  -2.956  39.505  1.00  1.00      B
ATOM   5675  CZ  ARG B 365      -3.984  -3.707  40.434  1.00  1.00      B
ATOM   5676  NH1 ARG B 365      -2.678  -3.913  40.385  1.00  1.00      B
ATOM   5677  NH2 ARG B 365      -4.698  -4.247  41.418  1.00  1.00      B
ATOM   5678  C   ARG B 365      -4.459  -2.492  34.648  1.00  1.00      B
ATOM   5679  O   ARG B 365      -5.449  -1.961  34.150  1.00  1.00      B
ATOM   5680  N   LEU B 366      -4.267  -3.801  34.609  1.00 41.59      B
ATOM   5681  CA  LEU B 366      -5.272  -4.665  33.996  1.00 44.25      B
ATOM   5682  CB  LEU B 366      -4.615  -5.908  33.366  1.00 45.24      B
ATOM   5683  CG  LEU B 366      -3.640  -5.701  32.202  1.00 45.46      B
ATOM   5684  CD1 LEU B 366      -4.331  -5.029  31.031  1.00 47.09      B
ATOM   5685  CD2 LEU B 366      -2.489  -4.856  32.678  1.00 46.71      B
ATOM   5686  C   LEU B 366      -6.263  -5.080  35.092  1.00 45.55      B
ATOM   5687  O   LEU B 366      -6.424  -6.296  35.333  1.00 46.32      B
ATOM   5688  OXT LEU B 366      -6.868  -4.169  35.704  1.00 46.33      B
ATOM   5689  CB  ARG C  10       -5.663   0.205  32.737  0.76  1.00      C
ATOM   5690  CG  ARG C  10       -7.073  -0.397  32.771  0.76  1.00      C
ATOM   5691  CD  ARG C  10       -7.748  -0.383  31.408  0.76  1.00      C
ATOM   5692  NE  ARG C  10       -8.728  -1.462  31.268  0.76  1.00      C
ATOM   5693  CZ  ARG C  10       -9.992  -1.301  30.875  0.76  1.00      C
ATOM   5694  NH1 ARG C  10      -10.464  -0.093  30.582  0.76  1.00      C
ATOM   5695  NH2 ARG C  10      -10.779  -2.365  30.749  0.76  1.00      C
ATOM   5696  C   ARG C  10       -4.106   2.152  32.497  0.76  1.00      C
ATOM   5697  O   ARG C  10       -3.278   1.863  33.369  0.76  1.00      C
ATOM   5698  N   ARG C  10       -6.417   2.186  31.464  0.76  1.00      C
ATOM   5699  CA  ARG C  10       -5.587   1.727  32.625  0.76  1.00      C
ATOM   5700  N   GLN C  11       -3.805   2.853  31.408  0.76  1.00      C
ATOM   5701  CA  GLN C  11       -2.458   3.321  31.094  0.76  1.00      C
ATOM   5702  CB  GLN C  11       -2.423   3.866  29.662  0.76  1.00      C
ATOM   5703  CG  GLN C  11       -1.047   4.361  29.231  0.76  1.00      C
ATOM   5704  CD  GLN C  11       -0.039   3.245  29.174  0.76  1.00      C
ATOM   5705  OE1 GLN C  11       -0.263   2.232  28.494  0.76  1.00      C
ATOM   5706  NE2 GLN C  11        1.082   3.415  29.876  0.76  1.00      C
ATOM   5707  C   GLN C  11       -1.895   4.396  32.038  0.76  1.00      C
```

```
ATOM   5708  O   GLN C  11      -2.494   5.467  32.217  0.76  1.00      C
ATOM   5709  N   LEU C  12      -0.732   4.111  32.618  0.76  1.00      C
ATOM   5710  CA  LEU C  12      -0.065   5.046  33.519  0.76  1.00      C
ATOM   5711  CB  LEU C  12       0.754   4.277  34.561  0.76  1.00      C
ATOM   5712  CG  LEU C  12      -0.036   3.305  35.450  0.76  1.00      C
ATOM   5713  CD1 LEU C  12       0.907   2.681  36.468  0.76  1.00      C
ATOM   5714  CD2 LEU C  12      -1.184   4.040  36.153  0.76  1.00      C
ATOM   5715  C   LEU C  12       0.845   5.948  32.680  0.76  1.00      C
ATOM   5716  O   LEU C  12       1.111   5.653  31.510  0.76  1.00      C
ATOM   5717  N   VAL C  13       1.317   7.044  33.273  0.76  1.00      C
ATOM   5718  CA  VAL C  13       2.166   7.987  32.543  0.76  1.00      C
ATOM   5719  CB  VAL C  13       1.473   9.371  32.386  0.76  1.00      C
ATOM   5720  CG1 VAL C  13       0.217   9.239  31.523  0.76  1.00      C
ATOM   5721  CG2 VAL C  13       1.113   9.929  33.750  0.76  1.00      C
ATOM   5722  C   VAL C  13       3.542   8.211  33.174  0.76  1.00      C
ATOM   5723  O   VAL C  13       3.740   8.050  34.381  0.76  1.00      C
ATOM   5724  N   LEU C  14       4.498   8.596  32.339  0.76  1.00      C
ATOM   5725  CA  LEU C  14       5.860   8.846  32.803  0.76  1.00      C
ATOM   5726  CB  LEU C  14       6.836   8.819  31.619  0.76  1.00      C
ATOM   5727  CG  LEU C  14       6.972   7.481  30.889  0.76  1.00      C
ATOM   5728  CD1 LEU C  14       7.666   7.705  29.557  0.76  1.00      C
ATOM   5729  CD2 LEU C  14       7.744   6.495  31.769  0.76  1.00      C
ATOM   5730  C   LEU C  14       6.010  10.186  33.517  0.76  1.00      C
ATOM   5731  O   LEU C  14       5.238  11.126  33.284  0.76  1.00      C
ATOM   5732  N   GLY C  15       7.000  10.263  34.396  0.76  1.00      C
ATOM   5733  CA  GLY C  15       7.264  11.510  35.090  0.76  1.00      C
ATOM   5734  C   GLY C  15       8.263  12.275  34.234  0.76  1.00      C
ATOM   5735  O   GLY C  15       9.472  12.210  34.462  0.76  1.00      C
ATOM   5736  N   LEU C  16       7.750  12.995  33.241  0.76  1.00      C
ATOM   5737  CA  LEU C  16       8.576  13.756  32.306  0.76  1.00      C
ATOM   5738  CB  LEU C  16       7.732  14.157  31.094  0.76  1.00      C
ATOM   5739  CG  LEU C  16       7.258  12.955  30.269  0.76  1.00      C
ATOM   5740  CD1 LEU C  16       6.303  13.411  29.171  0.76  1.00      C
ATOM   5741  CD2 LEU C  16       8.467  12.233  29.690  0.76  1.00      C
ATOM   5742  C   LEU C  16       9.263  14.982  32.898  0.76  1.00      C
ATOM   5743  O   LEU C  16      10.182  15.515  32.231  0.76  1.00      C
ATOM   5744  OXT LEU C  16       8.870  15.398  34.009  0.76  1.00      C
END
```

wherein atoms 4045 to 5688 represent the peptide binding site and atoms 5689 to 5748 represent the peptide.

The atomic coordinates are represented in protein data bank (pdb) format. Such a format is well known to the man skilled in the art.

According to another embodiment, the invention relates to a method to purify the processivity clamp factor of DNA polymerase, in particular the β subunit of DNA polymerase III of *Escherichia coli*, comprising the following steps:

- elution of a solution containing the processivity clamp factor of DNA polymerase, in particular the β subunit of DNA polymerase III of *Escherichia coli*, through a cation exchange column, in particular a SP sepharose column;
- elution of a solution containing the processivity clamp factor of DNA polymerase, in particular the β subunit of DNA polymerase III of *Escherichia coli*, in particular as obtained by the preceding step, through an anion exchange column, in particular a Mono Q column;
- elution of a solution containing the processivity clamp factor of DNA polymerase, in particular the β subunit of DNA polymerase III of

Escherichia coli, in particular as obtained by the preceding step, through a cation exchange column, in particular a Mono S column.

The expression "purify" relates to the process of separating a protein of interest from substantially all the other components of a solution containing said protein of interest, such as a bacterial extract.

Assessment of the purity of the protein of interest can be carried out by methods well known to the man skilled in the art, such as polyacrylamide gel electrophoresis analysis and Coomassie Blue staining or other type of protein staining (e.g. silver staining), mass spectrometry, protein sequencing, HPLC (high performance liquid chromatography). Quantification can be measured by absorbance spectroscopy, Bradford colorimetric assay, or protein sequencing.

The SP sepharose column, Mono Q column and Mono S column are obtained from Pharmacia (Uppsala, Sweden).

Alternatively, columns carrying ion exchange groups with properties similar to those of the SP sepharose column, Mono Q column and Mono S column can also be used.

The above mentioned column can be used with a FPLC system (Pharmacia), and possesses a high protein binding capacity. Advantageously, the SP sepharose column is used during the initial steps of the purification process because it is usually not clogged by dirty samples. The Mono Q and Mono S column are used during the last steps of the purification process, they are highly resolutive columns, but they are easily clogged by dirty samples.

The invention also relates to a method to obtain a protein crystal as defined above, comprising the following steps:

- mixing a solution of processivity clamp factor of DNA polymerase, with a solution of a peptide of about 3 to about 30 amino acids, in particular of about 16 amino acids, said peptide comprising all or part of the processivity clamp factor binding sequence of a processivity clamp factor interacting protein, such as prokaryotic Pol I, Pol II, Pol III, Pol IV, Pol V, MutS, ligase I, α subunit of DNA polymerase, UmuD or UmuD', or eukaryotic pol ε, pol δ, pol η, pol ι, pol κ, and with a solution of MES pH 6.0 0.2 M, $CaCl_2$ 0.2 M, PEG 400 60%, to obtain a crystallisation drop,

- letting the crystallisation drop concentrate against a solution of MES pH 6.0 0.1 M, $CaCl_2$ 0.1 M, PEG 400 30%, by vapour diffusion, to obtain a protein crystal.

The expression "vapour diffusion" refers to a crystallization method for macromolecules well known to the man skilled in the art, it is in particular described in "Crystallization of nucleic acids and proteins", pp. 130-145. A. Ducruix & R. Giegé eds., 1999, Oxford University Press.

MES refers to 2-(N-morpholino)-ethane sulfonic acid.

PEG 400 refers to polyethylene glycol 400.

Advantageously MES, PEG and $CaCl_2$ can be obtained from Hampton Research, (Laguna Niguel, USA).

The invention more particularly relates to a method to obtain a protein crystal as defined above, wherein the processivity clamp factor of DNA polymerase is the β subunit of DNA polymerase, in particular the β subunit of DNA polymerase III of *Escherichia coli*, in particular as purified according the abovementioned methods of purification, and the peptide has the following sequence:

VTLLDPQMERQLVLGL (SEQ ID NO: 1).

According to a preferred embodiment the β subunit of DNA polymerase III of *Escherichia coli* and the peptide of SEQ ID NO: 1 are mixed in a molar ratio of about 1:1 to about 1:3 in particular about 1: 1.5

According to another preferred embodiment the concentration of the β subunit of DNA polymerase III of *Escherichia coli* is from about 8 mg/ml to about 50 mg/ml , in particular about 34 mg/ml.

According to another preferred embodiment the concentration of the peptide of SEQ ID NO: 1 is from about 0.5 mg/ml to about 1.2 mg/ml, in particular about 1.1 mg/ml.

According to another embodiment, the invention relates to the use of the atomic coordinates as defined above, for the screening, the design or the modification of ligands of the processivity clamp factor of DNA polymerase, in particular of the β subunit of DNA polymerase III of *Escherichia coli*.

The expression "ligand" refers to a compound which is liable to bind to the processivity clamp factor of DNA polymerase.

The invention also relates to the use as defined above, for the screening, the design or the modification of ligands liable to be used for the preparation of pharmaceutical compositions useful for the treatment of bacterial diseases or diseases originating from DNA synthesis processes, such as fragile X syndrome, or proliferative disorders, such as cancers.

The expression "bacterial diseases" refers to diseases which are caused by bacterial influences, such as infections.

The expression "proliferative disorders" refers to disorders which are linked to abnormal cell multiplication, such as cancers.

The invention also relates to a method to screen ligands of the processivity clamp factor of DNA polymerase, said method comprising the step of assessing the interaction of tridimensional models of the ligands to screen with the structure of the β subunit of DNA polymerase as defined by the atomic coordinates as defined above, and in particular with the structure of the peptide binding site as defined by the atomic coordinates defined above, and more particularly with at least nine of the following amino acids: Leu 155, Thr 172, Gly 174, His 175, Arg 176, Leu 177, Pro 242, Arg 246, Val 247, Phe 278, Asn 320, Tyr 323, Val 344, Ser 346, Val 360, Val 361, Met 362, Pro 363, Met 364, Arg 365, Leu 366.

Assessing the interaction can be done by methods such as molecular dynamics, energy calculation, continuum electrostatics, semi-empirical free energy functions and other related methods well known to the man skilled in the art. Several packages and softwares are available for these purposes such as CHARM, UHBD, or SYBILL.

The invention more particularly relates to a method as defined above, to screen ligands liable to be used for the preparation of pharmaceutical compositions useful for the treatment of bacterial diseases or diseases originating from DNA synthesis processes, such as fragile X syndrome, or proliferative disorders, such as cancers.

The invention also relates to a method to design or to modify compounds liable to bind to the processivity clamp factor of DNA polymerase, said method comprising the step of designing or modifying a compound, so that the tridimensional model of said compound is liable to interact with the structure of the β subunit of DNA polymerase as defined by the atomic coordinates as defined above, and in particular with the structure of the peptide binding site as defined by the atomic coordinates as defined above, and more particularly with at least nine of the following amino acids: Leu 155, Thr 172, Gly

174, His 175, Arg 176, Leu 177, Pro 242, Arg 246, Val 247, Phe 278, Asn 320, Tyr 323, Val 344, Ser 346, Val 360, Val 361, Met 362, Pro 363, Met 364, Arg 365, Leu 366.

The invention more particularly relates to a method as defined above, to design or to modify ligands liable to be used for the preparation of pharmaceutical compositions useful for the treatment of bacterial diseases or diseases originating from DNA synthesis processes, such as fragile X syndrome, or proliferative disorders, such as cancers.

According to another embodiment, the invention relates to a peptide of the following sequence:

VTLLDPQMERQLVLGL (SEQ ID NO: 1).

According to a preferred embodiment, said peptide comprises non-hydrolysable bonds between amino-acids and/or non-amide bonds between amino-acids.

The invention also relates to a pharmaceutical composition comprising as active substance the peptide of SEQ ID NO: 1, in association with a pharmaceutically acceptable carrier.

Examples of pharmaceutically acceptable carrier are well known to the man skilled in the art.

According to a preferred embodiment, said peptide comprises non-hydrolysable bonds between amino-acids and/or non-amide bonds between amino-acids.

According to another embodiment the invention relates to the use of the peptide of SEQ ID NO: 1, as an anti-bacterial compound.

The expression "anti-bacterial compound" refers to a compound which has bactericidal or bacteriostatic properties, such as an antibiotic.

According to a preferred embodiment, said peptide comprises non-hydrolysable bonds between amino-acids and/or non-amide bonds between amino-acids.

The invention more particularly relates to the use of the peptide of SEQ ID NO: 1 for the manufacture of a medicament for the treatment of bacterial diseases or diseases originating from DNA synthesis processes, such as fragile X syndrome, or proliferative disorders, such as cancers.

According to another embodiment the invention relates to a method to test *in vitro* the inhibitory effect of compounds on the processivity clamp factor-dependant activity of DNA polymerase, in particular of Pol IV DNA polymerase of *Escherichia coli*, or of the α subunit of Pol III DNA polymerase of *Escherichia coli*, comprising the following steps:

- adding to assay solutions comprising a labelled nucleotidic primer, a template DNA, and DNA polymerase, in particular Pol IV DNA polymerase of *Escherichia coli*, or the α subunit of Pol III DNA polymerase of *Escherichia coli*, a compound to test at a given concentration for each assay solution, in the presence or the absence of the processivity clamp factor of DNA polymerase, in particular the β subunit of DNA polymerase in particular the β subunit of DNA polymerase III of *Escherichia coli*,

- electrophoretically migrating the abovementioned assay solutions,

- comparing the migration pattern of each assay solutions in the presence or the absence of the processivity clamp factor of DNA polymerase, in particular the β subunit of DNA polymerase, in particular the β subunit of DNA polymerase III of *Escherichia coli*.

According to a preferred embodiment of the above defined *in vitro* test method, the assay solutions also comprise a clamp loader, in particular the γ complex of *E. coli*, adenosine triphosphate (ATP), the divalent cation Mg^{2+} and single strand binding protein (SSB) of *E. coli*.

According to another preferred embodiment of the above mentioned *in vitro* test method, the compounds to be tested are such that their tridimensional models have been screened, modified or designed with respect to the structure of the β subunit of DNA polymerase, according to the corresponding above defined screening, modifying or designing methods.

The invention also relates to the use of the *in vitro* test method defined above, for the screening of compounds liable to be used for the preparation of pharmaceutical compositions useful for the treatment of bacterial diseases or diseases originating from DNA synthesis processes, such as fragile X syndrome, or proliferative disorders, such as cancers.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1

Figure 1 represents the atomic coordinates in protein databank (pdb) format of the crystallographic structure of the complex between *Escherichia coli* β subunit of DNA polymerase III and the 16 C-terminal residues of the β binding peptide of *E. coli* Pol IV DNA polymerase (P16)

Figure 2

Figure 2 represents a ribbon representation of the β subunit of DNA polymerase III of *E. coli* complexed with the P16 peptide (boxed) as obtained from the crystallographic structure of the complex.

Figure 3A, Figure 3B, Figure 3C and Figure 3D

Figure 3A and Figure 3B represent the inhibition of β dependant activity of Pol IV by the Pol IV β binding peptide, P16

Figure 3C and Figure 3D represent the inhibition of β dependant activity of Pol III α subunit by the Pol IV β binding peptide, P16.

Figure 3A represents the migration pattern of an electrophoresis gel. β free (lanes 1-4 and 9-12) or β loaded (lanes 5-8 and 13-16) labelled primer/template hybrids are incubated with increasing amounts of control peptide (CLIP) (lanes 1-8) or P16 peptide (lanes 9-16). Concentrations of peptides are as follows: 0 μM, lanes 1, 5, 9 and 13; 1 μM, lanes 2, 6, 10 and 14; 10 μM, lanes 3, 7, 11 and 15; 25 μM, lanes 4, 8, 12 and 16. This mixture is then submitted to the enzymatic activity of Pol IV (1.5 nM) in the presence of each four dNTPs for 1 minute at room temperature. Beside the overall increase in DNA synthesis activity, the β-dependent activity of the polymerase is characterised by the apparition of synthesis products longer than 12 nucleotides (*β* dependent synthesis), *β* independent synthesis is characterised by products shorter than 12 nucleotides. The broader band at the bottom of the gel corresponds to the primer.

Figure 3B represents the quantitative analysis of the relative amounts of each β-independent (incorporation of 1 up to 12 nucleotides) and β-dependent (12 and more nucleotides incorporation) activities observed in lanes 5-8 and 13-16. Black and white rectangles represent the ratio of β-dependent to β-independent polymerase activities

(vertical axis) in the presence of specified amounts of CLIP and P16 peptides (horizontal axis), respectively. Decrease in this ratio value actually indicates a specific inhibition of the β-dependent polymerase activity.

Figure 3C and 3D respectively correspond to the same experiments than those represented in Figure 3A and 3B, except that the polymerase used is the purified α subunit of Pol III (6 nM).

Figure 4

Figure 4 represents the growth rate of *E. coli* transformed by IPTG inducible plasmids expressing either the wild type Pol IV (pWp4) (triangles) or the Pol IVD5 mutant of Pol IV lacking the 5 C-terminal amino-acids (pD5p4) (squares, dotted line) in the presence of IPTG. The vertical axis represents the OD at 600 nm and the horizontal axis the time in minutes.

Figure 5A and Figure 5B

Figure 5A represents the growth rate of independent *E. coli* clones harbouring the P403FL vector in the absence (diamonds, triangles, crosses) or the presence (squares, dashes, circles) of 0.1 mM IPTG.

Figure 5B represents the growth rate of independent *E. coli* clones harbouring the P403D5 vector in the absence (diamonds, triangles, crosses) or the presence (squares, dashes, circles) of 0.1 mM IPTG.

The vertical axis represents the O.D. at 600 nm and the horizontal axis represents the time (in minutes).

Figure 6

Figure 6 represents Petri dishes containing an agarose-based nutritive medium supplemented with 0.05 mM IPTG and plated with *E. coli* cells harbouring P403FL (top) or with *E. coli* cells harbouring P403D5 (bottom).

EXAMPLES

EXAMPLE 1

Crystallographic study of the *Escherichia coli* β sliding clamp complexed with the β binding peptide of Pol IV DNA Polymerase of *E. coli*.

1. β binding peptide synthesis and purification

The 16-mer peptide sequence VTLLDPQMERQLVLGL (P16) (SEQ ID NO: 1), representing the 16 last residues of Pol IV DNA polymerase of *E. coli*, was obtained purified from Neosystem (Illkirch, France) and the 22-mer control peptide RPVKVTPNGAEDESAEAFPLEF (CLIP) (SEQ ID NO: 2) was a gift from Dr J.P. Briand (Strasbourg, France). P16 was resuspended at 1.1 mg/ml in a buffer containing Tris HCl 20 mM, pH 7.5, 5 mM EDTA, 20% glycerol, and kept at –80°C. CLIP was resuspended in 20 mM $NaHCO_3$ buffer, pH 9, at concentrations of 250, 100 and 10 pmoles/µl

2. β protein purification

The *dnaN* gene encoding *E. coli* β sliding clamp (hereafter referred to as β protein) was cloned into the pET15b plasmid (Invitrogen). The β protein was expressed in a transformed *E. coli* BL21(DE3)pLysS/(pET15b-*dnaN*) and was purified as described (Johanson *et al.*, 1986) with the following modifications. A SP Sepharose column (Pharmacia, Upsalla, Sweden) was used instead of the SP Sephadex column. A Mono Q column (Pharmacia, Upsalla, Sweden) followed by a Mono S column (Pharmacia, Upsalla, Sweden) were performed after the SP Sepharose column step. The β protein was purified to >99% purity, as judged by Coomassie gel analysis, and concentrated using Centriplus YM-30 concentrators (Amicon) to 34.2 mg/ml in a buffer containing 20 mM Tris-HCl pH 7.5, 0.5 mM EDTA and 20% (v/v) glycerol, as determined by Bradford assay, using BSA as a standard.

3. Crystalization conditions

Drops were obtained by mixing 0.92 µL of β protein at 34.2 mg/ml (775 pmoles) with 1.89 µl of P16 at 1.1 mg/ml (1136 pmoles) and 1 µl of 2X reservoir solution. Reservoir solution contains 0.1 M MES pH 6.0, 0.1M $CaCl_2$ and 30% PEG 400 (Hampton Research, Laguna Niguel, CA, USA). The peptide/β monomer molar ratio was 1.46. Co-crystals were

grown by vapour diffusion in hanging drops at 20°C. They typically grew within three days and reached 200 x 100 x 40 μm^3. Cristals were mounted in loops (Hampton Research, Laguna Niguel, CA, USA), frozen in liquid ethane and kept in liquid nitrogen before collection of crystallographic data.

4. Data collection and structure determination

Diffraction data were collected at beam line ID 14-EH4 (ESRF, Grenoble, France). The data were integrated with DENZO and normalized with SCALEPACK (Z. Otwinowski and W. Minor "Processing of X-ray Diffraction Data Collected in Oscillation Mode", Methods in Enzymology, Volume 276; Macromolecular Crystallography, part A, p. 307-326, 1997, C.W. Carter, Jr. and R.M. Sweet, Eds., Academic Press (New York)). The structure was solved by molecular replacement with MOLREP (CCP4, COLLABORATIVE COMPUTATIONAL PROJECT, NUMBER 4. (1994) "The CCP4 Suite: Programs for Protein Crystallography". Acta Cryst. D50, 760-763.), using the known β protein structure as a search model (Kong *et al.*, 1992). The peptide was built with the graphics program O (Copyright 1990 by Alwyn Jones, DatOno AB, Blueberry Hill, S-75591 Uppsala, Sweden) and the model was refined with O and CNS (Brunger *et al.*, 1998) (Copyright © 1997-2001 Yale University).

The results are summarized in following Table 1:

Data collection	
Space group	P1
Cell parameters	a=41.23 Å; b=65.22 Å; c=73.38 Å; α=73.11°; β=85.58°; γ=85.80°
X-ray source	ID14eh4
Wavelenght (Å)	0.93922
Asymetric unit	1 dimer
Resolution (Å)	1.65
Number of observations	
Unique	85999
Total	231008
Completeness (%)	96.7 (95.4)[a]
Rsym	0.051 (0.254)[a]
Mean I/σ	15.5 (4.3)[a]

Refinement	
Resolution range (Å)	500-1.65
R-factor, reflections	20.87, 80566
Rfree, reflexions	23.71, 4226
Number of atoms	
Protein	5744
Water	443
R.m.s deviation	
Bond angles (°)	1.59
Bond lenghts (Å)	0.013
Average atomic B-value (Å²)	
Protein	
β	22.8
Peptide	29.7
Water	29.1
Ramachandran plot [b] (%)	
residues in core,	92.4
allowed,	6.9
generously allowed regions	0.8

[a] Number in parentheses is for the last shell (1.71-1.65)
[b] Statistics from *PROCHECK* (Laskowski *et al.*, 1993)

Table 1: Crystal structure data and refinement statistics

The results obtained indicate that the crystal is triclinic, with cell dimensions a= 41.23 Å, b= 65.22Å, c= 73.38 Å, α= 73.11°, β= 85.58°, γ= 85.79°. These cell parameters lead to a quite usual value of 2.36 Å^3/Dalton for two molecules (i.e. one ring) per asymmetric unit. The present structure was solved by molecular replacement with the program MOLREP and was refined up to 1.65 Å resolution, which represents an important improvement in comparison to the 2.5 Å resolution obtained for the structure published previously (Kong *et al.*, 1992). The atomic coordinates of the structure solved by the Inventors are given in **Figure 1** in pdb format. The superposition of the present structure onto the previous one yields an overall rmsd of 1.22 Å for the Cα chain, which indicates that both structures are very similar, although numerous side chains and several mobile loops were rebuilt and a better description of the solvent was achieved. A more sensible superposition, systematically downweighting too distant residues (as those in the rebuilt loops), yields a weighted rmsd of 0.78 Å, which is more significant than the former value.

A density related to the presence of the peptide could be located after several rounds of refinement in a "simulated annealing composite omit map" (Brunger *et al.*, 1998). The seven C-terminal residues of the P16 peptide, $R_{10}Q_{11}L_{12}V_{13}L_{14}G_{15}L_{16}$, encompassing the

β binding sequence were built into the density map (**Figure 2**). This map extended slightly toward the N-terminus of the peptide but rapidly faded, so that the Q_{11} residue was still easily seen while the R_{10} was built in a poor density region. The rest of the peptide, probably disordered, was not visible. The seven C-terminal amino acids of the P16 peptide bind onto the β surface within two distinct but adjacent domains: one deep crevice, located between sub-domains 2 and 3 (area 1), and a second area which extends over the third β subdomain, close to the C-terminal extremity of the β chain (area 2) (**Figure 2**).

In the first area (area 1) of the peptide P16 binding site, two β strands of the clamp ($\beta^{4'}$ of domain 2 and $\beta^{8''}$ of domain 3) align. Some of their residues (L177 and V360, respectively), along with residues of the subdomain connecting loop (P242 and V247), form a hydrophobic pocket at the surface of the β monomer. The P16 residues L16 and L14 bind in this crevice. The hydrophobic nature of the interactions is revealed by the removal, upon peptide binding, of water molecules nested inside the free pocket. However, L14 and L16 are also involved in interactions with other adjacent residues like L155, T172, H175, R176, S346 and M362 (**Table 2**). The residue G15 has no interaction with any residues of the pocket and serves as a connector between L14 and L16. Consequently, the L16 residue which, according to the pentapeptidic consensus motif ($Q_1L_2(SD)_3L_4F_5$) (Dalrymple *et al.*, 2001), was not considered to belong to the β-binding sequence, actually fully participates to the interaction.

In the second binding area (area 2), the four other P16 residues, V13, L12, Q11 and R10 establish mostly hydrophobic interactions with residues H175, N320, Y323, V344, M362, P363 and M364 of the β monomer (**Table 2**). Among the four P16 residues located within this region, the Q residue is highly conserved within the binding motifs of the various β ligands, to the same extent as residues that bind into the hydrophobic crevice (L14 and L16) (Dalrymple *et al.*, 2001). Particularly, it forms interactions, directly or mediated by two water molecules with β residues M362 and E320. These contacts might prime the binding of the peptide with the β surface and facilitate the formation of interactions of the C-terminal residues within the hydrophobic pocket of area 1. Thus the peptide would be anchored on the β surface by two points located on each extremity of the binding sequence.

β residues	Interacting P16 residues
M364	R10,Q11,L12
P363	Q11, L12
M362	Q11,L12,V13,L14
V361	**L14**
V344	L12
Y323	Q11
N320	Q11
V360	L14
S346	L14
V247	L14,L16
P242	L16
L177	L14,**L16**
R176	L14
H175	Q11,L12,V13,L14
T172	**L14**,L16
L155	**L16**

Interactions between the β residues and the peptide P16 residues. All considered distances between β and peptide P16 residues are between 3 and 3.8 Å, except those (P16 residues in bold) between L155:L16, T172:L14, L177:L16 and V361:L14 which are larger than 4 Å.

Table 2

5. N-terminal sequencing of the protein

The cristal was recovered after data collection, washed several times in the well solution, and dissolved in 10 µl water. The proteins contained within the crystal were derivatized and sequenced by automated Edman's degradation using a PE Applied Biosystems 492 cLC Protein Sequencer allowing the identification and precise quantitative analysis of the amino acids released at each step of degradation.

6. Improvement of the P16-β clamp interaction

Preliminary *in silico* docking experiments carried out with modified versions of the P16 peptide suggest that its interaction with the β clamp could be strengthened by replacing Leu 12 and Leu 14 by aromatic amino acids, or by extending the lateral chain of Gln 11. Thus, these modifications show the way to designing new high affinity β clamp interaction inhibitors.

EXAMPLE 2

***In vitro* study of the β clamp-β binding peptide of Pol IV interaction by competition assays**

In order to ascertain the biological relevance of the P16 peptide-β clamp interaction observed in the crystallographic structure, an *in vitro* assay based on the activity of Pol IV DNA polymerase was designed. This assay relies on the observation that the *in vitro* activity of Pol IV is greatly enhanced by the presence of the β subunit loaded onto a primer/template DNA substrate (Wagner *et al.*, 2000) (**Figure 3A**, compare lanes 1 and 5 or 9 and 13), while the enzyme alone incorporates nucleotides in a distributive mode (Wagner *et al.*, 1999).

Briefly, P16 peptide and a control peptide (CLIP) were diluted in 20 mM $NaHCO_3$ at concentrations of 250, 100 and 10 pmol/µl. 5' end radiolabelling, purification and annealing of synthetic primers were performed as previously described (Wagner *et al.*, 1999). The 30/90 nucleotide synthetic construct (Wagner *et al.*, 2000) was obtained by annealing the 30 nucleotide primer (5'GTAAAACGACGGCCAGTGCCAAGCTTAGTC) (SEQ ID NO : 3) with the 90 nucleotide template (5'CCATGATTACGAATTCAGTCATCACCGGCGC CACAGACTAAGCTTGGCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACC CTGG) (SEQ ID NO : 4) to form a double stranded structure with 5' and 3' single stranded DNA overhangs of 25 and 35 nucleotides, respectively.

All replication experiments (10 µl final volume) were carried out in buffer E (40 mM HEPES pH 7.5, 80 mM potassium glutamate, 160 µg/ml BSA, 16 % glycerol, 0.016 % NP40, 8 mM DTT). The 30/90 nucleotide hybrid was first incubated with single strand binding proteins (SSB; Sigma; 90 nM final concentration) in the presence of ATP (200 µM) and $MgCl_2$ (7.5 mM) at 37°C for 10 min. When specified, the γ complex (1 nM final concentration) (gift from Dr. C. S. McHenry, Denver, USA), and the β clamp (5 nM as dimer final concentration) were added at that stage, and incubation was carried out at 37°C for 10 min. Then, 7 µl of the mixture was added to 1 µl of either 20 mM $NaHCO_3$ or 1 µl of peptide solution (1, 10 or 25 µM final concentration), incubated 20 min. at room temperature and further 2 hours at 4°C. 1 µl of polymerase was then added (1.5 nM of Pol IV or 6 nM of α subunit (gift from Dr. H. Maki, Nara, Japan) final concentrations), incubated 5 min. at room temperature and finally, the whole reaction was mixed with 1 µl of a dNTPs solution (200 µM each dNTP final concentration) and let to react for 1 min. at room temperature. Reactions were quenched by the addition of 20 µl of 95 % formamide/dyes solution containing 7.5 mM

EDTA, heat-denatured and analysed by chromatography on 12 % denaturing polyacrylamide gels. Radiolabelled products were visualised and quantified using a PhosphorImager 445 SI (Molecular Dynamics) and the ImageQuant software.

As shown in **Figure 3A** and **Figure 3B**, increasing amounts of P16 inhibits the β-dependent activity of Pol IV (lane 13 to 16). At the highest P16 concentration tested (25 μM), the β-dependent Pol IV activity is decreased by a factor around 30, as indicated on the graphic. On the other hand, the control peptide (CLIP) has no effect on this activity even at the highest concentration tested (**Figure 3A**, lane 8). Also, neither P16 nor CLIP peptides do affect the intrinsic activity of Pol IV characterised by the distributive incorporation of one to up to 12 nucleotides (**Figure 3A**, lanes 1-4, 9-12, **Figure 3B**). Thus P16 specifically inhibits the β-Pol IV DNA polymerase interaction in solution, which demonstrate that the site we identified actually corresponds to the Pol IV DNA polymerase binding site on β.

The polymerase activity of the α subunit of the replicative DNA Polymerase III of *E. coli* is greatly enhanced by its interaction with the β clamp (Marians *et al.*, 1998) (**Figure 3C**, compare lanes 1 and 5 or 9 and 13), and the putative β binding peptide of the α subunit has been identified through bioinformatics analysis (Dalrymple *et al.*, 2001) and is a variant of the pentapeptide consensus motif. In order to determine if the replicative DNA polymerase interact with the β monomer within the same site than Pol IV, the ability of P16 peptide to inhibit the β-dependent activity of the α subunit was tested. The dose dependent inhibition of the α subunit β-dependent activity (**Figure 3C**, lane 13 to 16, **Figure 3D**) strongly suggest that this is the case. To achieve a high level of inhibition, the concentration of P16 peptide should exceed the polymerase concentration by a factor of 4 to 16.10^3. The need for such a high excess of peptide may reflect a higher affinity of the whole protein for the DNA-β substrate, mediated by other polymerase-β and/or polymerase-DNA interactions, but also a high entropic factor of the free peptide as opposed to the same fragment folded in the whole protein. Therefore, the lower peptide affinity would result from a lower kinetic constant k_{on}, and not from an increased k_{off}. Overall, this biochemical analysis indicates that (i) the P16-β structure we solved is of biological significance as indicated by the competitive inhibition of the β dependent activity of Pol IV DNA polymerase by peptide P16 and (ii) that peptide P16 also competes with and inhibits the β dependent activity of the α subunit of the DNA Polymerase III of *E. coli* which suggests that (iii) if not identical, the Pol IV and α subunit interaction sites on β subunit overlap.

EXAMPLE 3

***In vivo* study of the inhibition of bacterial growth by the β binding peptide of Pol IV**

Plasmids bearing either the wild type Pol IV (pWp4) or the Pol IV mutant deleted for the 5 last C-terminal residues (pD5p4) coding sequences under the IPTG inducible *lac* promoter were transformed into recipient *E. coli* cells (BL21(DE3, pLys)). These transformed cells were then allowed to grow in LB medium at 37°C with aeration and without or with (**Figure 4**) addition of the protein expression inducer IPTG (0.1 mM). Growth rates were monitored by measuring the optical density of the cultures (OD 600 nm) at different time points.

The growth rates of both cultures without artificial protein expression were identical whether the cells contain the wild type Pol IV expression plasmid (pWp4) or the Pol IVD5 mutant (pD5p4). On the other hand, when protein expression was induced by the adjunction of low IPTG concentration in the culture medium (**Figure 4**), a clear growth inhibition was observed for the culture expressing the wild type Pol IV protein compared to the one expressing the mutant protein. As the mutant protein (expressed from pD5p4) lacks essential amino acids for the interaction with the β-clamp, the observed cytotoxicity may be rationalised by the fact that the wild type Pol IV protein interacts with the β clamp and, because of its relative high concentration, interfere and/or compete with the binding of the replicative DNA polymerase, thereby inhibiting chromosome replication and culture growth.

In other words, these preliminary results indicate that site-specific β binding molecules (such as the Pol IV β binding motif) may serve as antimicrobial agents.

EXAMPLE 4

***In vivo* study of the inhibition of bacterial growth by the β binding peptide of Pol IV**

A DNA sequence encoding a catalytically inactive version of DNA polymerase IV of *E. coli* has been cloned into a vector to form P403FL which enable the IPTG inducible expression of the corresponding inactive enzyme. Similarly, a DNA sequence encoding the catalytically inactive version of DNA polymerase IV of *E. coli* depleted of the 5 last C-terminal residues (which are essential residues for the interaction with the β clamp) has been cloned into the same IPTG inducible vector to form P403D5.

Three independently isolated clones of *E. coli* containing either P403FL or P403D5 were cultured in a selective medium until an optical density (O.D.) of 0.2 at 600 nm was reached, 15 ml of a selective medium containing 0 or 0.1 mM IPTG were then inoculated with a quantity corresponding to 0.02 O.D. unit of the culture and bacterial growth was followed by the measure of the optical density at 600 nm during 5 hours.

The results indicate that in the absence of IPTG the three cultures of the independent clones carrying P403FL grow normally, however, in the presence of 0.1 mM IPTG the growth of these clones is completely halted (**Figure 5A**). Conversely, the three independent clones carrying P403D5 grow normally, irrespective of the presence or not of IPTG (**Figure 5B**).

Furthermore, about 1000 *E. coli* cells harbouring either P403FL or P403D5 were plated on nutritive agarose dishes containing 0.05 mM IPTG. The results shown in **Figure 6** indicate that, whereas essentially no P403FL carrying cells are growing, essentially all P403D5 carrying cells are growing.

As in **Example 3**, those results confirm that site-specific β binding molecules (such as the Pol IV β binding motif) may serve as antimicrobial agents.

REFERENCES

Becherel O. *et al.* (2002) "Pivotal role of the b-clamp in translesion DNA synthesis and mutagenesis in *E. coli* cells." DNA Repair **68**: 1_6.

Becherel O. and Fuchs R. P. (2001) "Mechanism of DNA polymerase II-mediated frameshift mutagenesis." Proc Natl Acad Sci U S A **98**(15): 8566-71.

Bridges B. A. and Woodgate R. (1985). "The two-step model of bacterial UV mutagenesis." Mutat Res **150**(1-2): 133-9.

Bruck, I. and O'Donnell M. (2001). "The ring-type polymerase sliding clamp family." Genome Biol. **2**(1): 3001.1-3000.3.

Brunger A. T. *et al.*, (1998) "Crystallography and NMR system (CNS) a new software system for macromolecular structure determination." Acta Cryst. **D54**: 905-921.

Cordonnier A. M. *et al.* (1999) "Impaired translesion synthesis in xeroderma pigmentosum variant extracts." Mol Cell Biol **19**(3): 2206-11.

Dalrymple B. P. *et al.* (2001) "A universal protein-protein interaction motif in the eubacterial DNA replication and repair systems." Proc Natl Acad Sci U S A **98**(20): 11627-32.

Friedberg *et al.* (2000a) "The many faces of DNA polymerases: strategies for mutagenesis or for mutational avoidance." Proc Natl Acad Sci USA **97**:5681-5683.

Friedberg *et al.* (2000b), "Specialized DNA polymerases, cellular survival, and the genesis of mutations." Science **296**:1627-1630.

Fuchs R. P. *et al.* (2001) "DNA polymerases II and V mediate respectively mutagenic (-2 frameshift) and error-free bypass of a single N-2-acetylaminofluorene adduct." Biochem Soc Trans **29**(Pt 2): 191-5.

Gulbis J. M. *et al.* (1996) "Structure of the C-terminal region of p21 complexed with human PCNA." Cell **87**:297-306.

Haracska L. *et al.* (2001a) "Physical and functional interactions of human DNA polymerase eta with PCNA." Mol Cell Biol **21**(21): 7199-206.

Haracska L. *et al.* (2001b) "Targeting of human DNA polymerase iota to the replication machinery via interaction with PCNA." Proc Natl Acad Sci U S A **98**(25): 14256-61.

Haracska L. *et al.* (2002) "Stimulation of DNA synthesis activity of human DNA polymerase kappa by PCNA." Mol Cell Biol **22**(3): 784-91.

Jeruzalmi D. *et al.* (2001) "Mechanism of processivity clamp opening by the delta subunit wrench of the clamp loader complex of *E. coli* DNA polymerase III." Cell 106(4): 417-28.

Kong, X. P. *et al.* (1992) "Three-dimensional structure of the beta subunit of *E. coli* DNA polymerase III holoenzyme: a sliding DNA clamp." Cell 69(3): 425-37.

Laskowski R A *et al.* (1993) "PROCHECK: a program to check the stereochemical quality of protein structures." J. Appl. Cryst. 26: 283-291.)

Lenne-Samuel N. *et al.* (2002) "The processivity factor beta controls DNA polymerase IV traffic during spontaneous mutagenesis and translesion synthesis *in vivo*." EMBO Rep 3(1): 45-9.

Marians K. J. *et al.* (1998) "Role of the core DNA polymerase III subunits at the replication fork." The Journal of biological Chemistry 273(4): 2452-2457.

Napolitano R. *et al.* (2000) "All three SOS-inducible DNA polymerases (Pol II, Pol IV and Pol V) are involved in induced mutagenesis." Embo J 19(22): 6259-65.

Ohmori H. *et al.* (2001). "The Y family of DNA polymerases", Mol Cell, 8: 7-8.

Reuven N. B. *et al.* (1999) "The mutagenesis protein UmuC is a DNA polymerase activated by UmuD', RecA, and SSB and is specialized for translesion replication." J Biol Chem. 274(45): 31763-6.

Shamoo Y. and Steitz T. A. (1999) "Building a replisome from interacting pieces: sliding clamp complexed to a peptide from DNA polymerase and a polymerase editing complex." Cell 99(2): 155-66.

Tang M. *et al.* (1999) "UmuD'(2)C is an error-prone DNA polymerase, *Escherichia coli* pol V." Proc Natl Acad Sci U S A 96(16): 8919-24.

Thompson J. D. *et al.* (2000) "DbClustal: rapid and reliable global multiple alignments of protein sequences detected by database searches." Nucleic Acids Research 28(15): 2919-2926.

Tsurimoto T. (1999) "PCNA binding proteins." in Front Biosci, vol 4:D 849-858.

Wagner J. *et al.* (2002) "Genetics of mutagenesis in *E. coli*: various combinations of translesion polymerases (Pol II, IV and V) deal with lesion/sequence context diversity." DNA Repair 1: 159-167.

Wagner J. *et al.* (2000) "The beta clamp targets DNA polymerase IV to DNA and strongly increases its processivity." EMBO Rep 1(6): 484-8.

Wagner J. *et al.* (1999) "The dinB gene encodes a novel *E. coli* DNA polymerase, DNA pol IV, involved in mutagenesis." Mol Cell 4(2): 281-6.

Willcox B. E. (1999) "TCR binding to peptide-MHC stabilizes a flexible recognition interface." Immunity 10: 357-65.

This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☑ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.
As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.